Week 3 Notes

Lectures 5 & 6

University of Massachusetts Amherst, CS389

1 Multilayer Perceptrons (lecture 5)

1.1 Non-linearities

Examples of non-linearity functions can be refered below

1.1.1 Sigmoid

$$\text{sig}(x) = \frac{1}{1 + e^{-x}} \tag{1.1}$$

The derivative is:

$$\frac{d}{dx}\text{sig}(x) = \text{sig}(x) \cdot (1 - \text{sig}(x)) \tag{1.2}$$

1.1.2 Tanh

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{1.3}$$

The derivative is:

$$\frac{d}{dx}\tanh(x) = 1 - \tanh^2(x) \tag{1.4}$$

1.1.3 ReLU

$$\text{ReLU}(x) = \max(0, x) \tag{1.5}$$

Derivative is:

$$\frac{d}{dx}\text{ReLU}(x) = \begin{cases} 0 & x < 0 \\ 1 & \text{otherwise} \end{cases} \tag{1.6}$$

2 Backpropagation (lecture 5)

> **Derivation:**
>
> The algebraic system define over operator $\star$, which is *closed* and *associative* is called SEMI-GROUP.

2.1 Code implementation (lecture 6)

References

[1] Cooper.